MAXIMIZING THE SCIENTIFIC RETURN FROM LUNAR SURFACE SCIENCE MISSIONS THROUGH DATA ARCHIVING AND SERVICES.

D. M. H. Baker[1], C. Acton[2], M. E. Banks[1], D. Crichton[2], L. Gaddis[3], T. McClanahan[1], T. Morgan[1], J. Padams[2], T. Stein[4], D. R. Williams[1]. [1]NASA Goddard Space Flight Center, Greenbelt, MD (david.m.hollibaughbaker@nasa.gov), [2]Jet Propulsion Laboratory, Pasadena, CA, [3]USGS, Astrogeology Science Center, Flagstaff, AZ, [4]Washington University in St. Louis, St. Louis, MO.

Introduction: The Planetary Data System (PDS) was established three decades ago to address community concerns that data returned by planetary science missions were in danger of being lost. With the expected large influx of lunar surface science data over the next decade, it is prudent to again underscore the importance of integrating proper data archiving into lunar mission data pipelines to preserve and maximize the scientific return of these missions. Here, we summarize current services and data that the PDS provides to the lunar science community and, drawing on the experience of the PDS and previous Apollo and planetary surface missions, we outline important considerations for the future archiving of and user access to lunar surface science data.

The PDS Data Services Initiative: In recent years, the PDS has developed and implemented a set of data archiving standards called "PDS4" to support the archiving of highly heterogeneous data from over 600 instruments [1]. PDS4 provides the basis for defining and capturing essential metadata (data or information that is given to describe or help you use other data), which then drives the development of data search, access, and discoverability. The International Planetary Data Alliance (IPDA) [2] endorsed PDS4 standards in 2012, which will allow enhanced discoverability of data across PDS and international archives. Future lunar science data will be archived using PDS4 standards, and will be a part of a planetary ecosystem that will involve the next generation of tools and services built on the foundation provided by PDS4 [1].

Current PDS Holdings Enable Lunar Science: A number of lunar datasets from past and active missions are currently archived in the PDS through the following Discipline Nodes: Atmospheres, Cartography and Imaging Sciences, Geosciences, Navigation and Ancillary Information Facility (NAIF), and Planetary Plasma Interactions. Visit https://pds.nasa.gov for more information on their holdings. Like all PDS compliant holdings, these data are peer-reviewed and follow strict archiving standards. Due mostly to the large volume of data collected by the Lunar Reconnaissance Orbiter (LRO) and its LRO camera (LROC) [3], lunar data comprise over 65% of the total PDS archive by volume [4]. These holdings provide the source material for numerous science investigations and lay the foundation for lunar site characterization for landed missions. The PDS archive of SPICE information also provides precision knowledge of the lunar-fixed reference frames (i.e., coordinate systems) for assisting scientists and engineers in planning and interpreting scientific observations from space-based instruments [5]. While the majority of the scientific data is currently archived using PDS3 standards, data is being migrated to PDS4 standards over the next several years.

The PDS also provides a suite of tools to enable search and access to archived lunar data products, including the widely used Orbital Data Explorer [6], Planetary Image Atlas [7], and tools and interfaces at the LROC Data Node [4]. As the lunar science data archive continues to grow and the PDS shifts toward more user-centric services, we anticipate a number of new and enhanced tools to enable data discovery between all lunar datasets archived in the PDS and international archives [1].

Lessons from Apollo: The Apollo missions serve as both a warning of archiving gone awry and a lesson for the proper handling of future lunar surface data. From 1969 to 1972, Apollo conducted a number of scientific experiments using orbiting platforms, lunar surface operations run by astronauts, and through astronaut-deployed packages that ran autonomously on the surface. Since Apollo occurred before formal archiving standards were conceived (i.e., well before the PDS), much of the data were not archived at the time they were collected, some were lost, and the data that were submitted for archiving lacked sufficient metadata [8,9]. Further, both storage media and formats were not standardized and many quickly became obsolete.

Thanks to recent efforts of the NASA Space Science Data Coordinated Archive (NSSDCA) through the Lunar Data Project, a number of Apollo science measurements have been recovered and digitally archived in the PDS [9]. However, even with this restoration, the improper original processing of the data and lack of original raw data has hampered the usability of some measurements [10].

As a result of the lack of a robust archiving plan, the Apollo data currently exists in disparate locations and formats with varying levels of completeness. Many digital connections to the rock and soil samples collected by Apollo astronauts also remain poor. This presents significant challenges to attempts to make Apollo mission data discoverable and understood in the context of the mission operations. As Feist et al. [11] demonstrated through the apollo17.org project, at the very least it is critically important that future lunar surface data share accurate time-tag and geospatial location metadata. This includes all images, videos, voice, samples, and scientific measurements recorded during the mission. For digital data archived with the PDS, all of this information can now be captured and described in the metadata using PDS4 standards.

Lessons from Surface Rovers: Since Apollo, the successful Mars rovers (Sojourner, Spirit, Opportunity, and Curiosity) [12,13] have offered an opportunity to identify and refine the needs for successful curation and tools for accessing complex planetary surface datasets. Like human surface operations, rovers have a diversity of measurements and complex operations that require carefully curated data and metadata. PDS tools such as the Geosciences Node's Analyst's Notebooks (AN) [6] integrate sequence information, engineering and science data, and documentation obtained from planetary surface missions into a web-based interface. The AN use the archived metadata and associated documentation, as well as information from non-archive sources, to facilitate a user's understanding of the data in the context of other observations, and the reasons for acquiring the data. As learned from Apollo, providing this context is particularly important for landed human missions, where the multitude of science measurements, samples, and human and robotic operational records are difficult to understand in isolation. Collaboration between the mission science team and data archiving team prior to launch and during surface operations greatly increases the value of the archives and tools.

The ANs have been extended for use for lunar datasets, including data restored from the Apollo missions and LCROSS [14]. It is envisioned that future lunar surface science missions would benefit from similar and more enhanced tools provided by the PDS.

Considerations for Data from Future Lunar Surface Science Missions: The wide use and robustness of current PDS holdings, and the lessons learned from Apollo, demonstrate that proper archiving of data is critical for the preservation of high quality and scientifically useful data. As the Apollo data have shown, data obtained fifty years ago are as relevant today as they were when they were collected. However, the data are only as useful to future generations as the robustness of our effort to preserve and document them. Based on this experience, we list a number of considerations for future lunar surface science missions:

1) Integrate PDS archive preparation and identify metadata needs early in the development of mission and data pipelines.

2) Archive raw, calibration, and higher-level data products with complete metadata to ensure greatest usability. Include data collected from instruments and supporting data (e.g., observation planning and targeting metadata) that provide context to measurements. For higher-order data, the PDS allows for archival of the popular GeoTIFF format for geospatial data [15] provided the file is also described using a PDS4 label.

3) Ensure that metadata common to all data products are captured to allow linkages between surface measurements. This includes common and accurate time and geospatial information.

4) Ensure common metadata linkages with rock and regolith samples and other field and operational documentation (verbal, written, visual).

5) Provide complete documentation on the instrument, data collection, calibration, and known issues.

6) Archive using data standards (i.e., PDS4) that enable linkages with other lunar and planetary datasets and that lay the foundation for user-centric search and discovery tools such as the Analyst's Notebooks.

Further, we stress that the PDS does not set the final archiving requirements for NASA missions. Mission program offices at NASA are the final authority for mission Data Archiving and Analysis Plans, including delivery schedules. It is therefore ultimately up to NASA and the missions to ensure that archiving is part of the plan from the beginning.

For science missions, the commitment to archiving has been strengthened under the recently signed "Science Mission Directorate's (SMD) Strategy for Data Management and Computing for Groundbreaking Science 2019-2024" [16], which notes that making data open and machine readable is the new default for government activities, and that the new open data policy will be implemented in Announcements of Opportunity (AOs), SALMON templates, and in Senior Reviews. The report recommends that "Missions should take a lifecycle approach in planning how their data will be managed for long-term curation after KDP-F". Archiving data in an approved permanent archive such as the PDS, satisfies that requirement.

References: [1] Crichton, D. et al. (2019) *4th Planetary Data Workshop*, Abstract #7105. [2] Crichton, D. et al. (2018) *AGU Fall Meeting*, Abstract # IN11D-0656. [3] Robinson, M. S. (2010) S*pace Sci. Rev.* 150, 81–124. [4] Estes, N. et al. (2019) 4th Planetary Data Workshop, Abstract #7082. [5] Acton, C. et al. (2018), PSS, 150, 9-12. [6] Stein, T. C. et al. (2015) *2nd Planetary Data Workshop*, Abstract #7019. [7] Grimes, K. et al. (2019) *4th Planetary Data Workshop*, Abstract #7093. [8] Williams, D. R. et al. (2015) *LPS XLVI*, Abstract #2101. [9] Williams, D.R. et al. (2018) *AGU Fall Meeting*, Abstract #P23E-3497. [10] Nakamura, Y. et al. (2015) *AGU Fall Meeting*, Abstract #IN33A-1788. [11] Feist, B. et al. (2018) *LPS XLIX*, Abstract #2681. [12] Crisp, J. A., et al. (2003) *JGR* 108(E12), 8061. [13] Grotzinger, J. et al. (2012) *Space Sci. Rev.* 170(5), 56. [14] Stein, T. C., et al. (2010) *LPS XLI*, Abstract #1414. [15] Hare, T., et al. (2020) *submitted to LPS LI*. [16] SMD Strategic Management Working Group (2019), Science Mission Directorate's (SMD) Strategy for Data Management and Computing for Groundbreaking Science 2019-2024", NASA Headquarters.